

PCT09

## RAW SEQUENCE LISTING

DATE: 01/14/2002

PATENT APPLICATION: US/09/856,327

TIME: 10:50:58

Input Set : A:\0230-0157P.ST25.txt

Output Set: N:\CRF3\01142002\I856327.raw

PS

3 <110> APPLICANT: TAKAKURA, Yoshimitsu et al.  
 5 <120> TITLE OF INVENTION: A NOVEL PROTEIN, A GENE CODING THEREFOR AND A METHOD OF  
 USING THE SAME

7 &lt;130&gt; FILE REFERENCE: 0230-0157P

9 &lt;140&gt; CURRENT APPLICATION NUMBER: 09/856,327

10 &lt;141&gt; CURRENT FILING DATE: 2001-05-21

12 &lt;160&gt; NUMBER OF SEQ ID NOS: 12

14 &lt;170&gt; SOFTWARE: PatentIn version 3.1

16 &lt;210&gt; SEQ ID NO: 1

17 &lt;211&gt; LENGTH: 2106

18 &lt;212&gt; TYPE: DNA

19 &lt;213&gt; ORGANISM: Lyophyllum shimeji

21 &lt;220&gt; FEATURE:

22 &lt;221&gt; NAME/KEY: CDS

23 &lt;222&gt; LOCATION: (8)..(1861)

24 &lt;223&gt; OTHER INFORMATION:

27 &lt;400&gt; SEQUENCE: 1

28	atcagcc	atg	tct	ctc	tca	acc	gag	cag	atg	cta	cgc	gac	tat	cca	cgg	49	
29		Met	Ser	Leu	Ser	Thr	Glu	Gln	Met	Leu	Arg	Asp	Tyr	Pro	Arg		
30		1				5					10						
32	tct	atg	caa	atc	aac	gga	cag	att	cct	aag	aac	gca	att	cac	gaa	aca	97
33	Ser	Met	Gln	Ile	Asn	Gly	Gln	Ile	Pro	Lys	Asn	Ala	Ile	His	Glu	Thr	
34	15				20				25					30			
36	tac	gga	aac	gac	gga	gtt	gat	gta	ttc	att	gca	gga	tct	gga	ccc	att	145
37	Tyr	Gly	Asn	Asp	Gly	Val	Asp	Val	Phe	Ile	Ala	Gly	Ser	Gly	Pro	Ile	
38					35				40					45			
40	gga	gcg	acg	tat	gca	aag	ctc	tgt	gtt	gaa	gct	ggt	cta	cgt	gtt	gtg	193
41	Gly	Ala	Thr	Tyr	Ala	Lys	Leu	Cys	Val	Glu	Ala	Gly	Leu	Arg	Val	Val	
42				50				55					60				
44	atg	gtc	gag	atc	gga	gct	gct	gat	agc	ttc	tac	gct	gtt	aat	gcc	gaa	241
45	Met	Val	Glu	Ile	Gly	Ala	Ala	Asp	Ser	Phe	Tyr	Ala	Val	Asn	Ala	Glu	
46		65				70				75							
48	gaa	gga	act	gca	gtt	ccc	tac	gtt	cct	ggc	tac	cac	aag	aag	aat	gaa	289
49	Glu	Gly	Thr	Ala	Val	Pro	Tyr	Val	Pro	Gly	Tyr	His	Lys	Lys	Asn	Glu	
50		80				85				90							
52	atc	gag	ttc	cag	aaa	gat	att	gac	cgc	ttc	gtc	aat	gta	atc	aag	gga	337
53	Ile	Glu	Phe	Gln	Lys	Asp	Ile	Asp	Arg	Phe	Val	Asn	Val	Ile	Lys	Gly	
54	95				100				105					110			
56	gcc	tta	caa	caa	gtc	tct	gtt	cct	gtc	aga	aac	cag	aac	gtg	cct	aca	385
57	Ala	Leu	Gln	Gln	Val	Ser	Val	Pro	Val	Arg	Asn	Gln	Asn	Val	Pro	Thr	
58				115				120					125				
60	ctt	gat	ccc	gga	gcc	tgg	agc	gcg	ccc	cct	gga	agt	tca	gcc	ata	tcg	433
61	Leu	Asp	Pro	Gly	Ala	Trp	Ser	Ala	Pro	Pro	Gly	Ser	Ser	Ala	Ile	Ser	
62				130				135					140				
64	aac	ggt	aaa	aat	cct	cac	cag	cgg	gaa	ttc	gag	aac	ttg	tct	gcg	gag	481
65	Asn	Gly	Lys	Asn	Pro	His	Gln	Arg	Glu	Phe	Glu	Asn	Leu	Ser	Ala	Glu	
66		145				150				155							
68	gcc	gta	acg	cgt	gga	gtc	ggc	ggc	atg	agt	acc	cac	tgg	acg	tgc	tcc	529

ENTERED

## RAW SEQUENCE LISTING

DATE: 01/14/2002

PATENT APPLICATION: US/09/856,327

TIME: 10:50:58

Input Set : A:\0230-0157P.ST25.txt

Output Set: N:\CRF3\01142002\I856327.raw

```

69 Ala Val Thr Arg Gly Val Gly Gly Met Ser Thr His Trp Thr Cys Ser
70      160      165      170
72 acg cca cgg att cat cca ccc atg gaa agt ctc ccg ggc atc ggc cgt      577
73 Thr Pro Arg Ile His Pro Pro Met Glu Ser Leu Pro Gly Ile Gly Arg
74 175      180      185      190
76 ccg aag ctc agt aac gac ccg gca gag gac gac aaa gag tgg aac gag      625
77 Pro Lys Leu Ser Asn Asp Pro Ala Glu Asp Asp Lys Glu Trp Asn Glu
78      195      200      205
80 ctt tat tcc gag gcc gag cgt ctc atc ggg act tcc acc aag gaa ttc      673
81 Leu Tyr Ser Glu Ala Glu Arg Leu Ile Gly Thr Ser Thr Lys Glu Phe
82      210      215      220
84 gac gag tca att cgg cac acc ctt gtt ctg cgc tct ttg caa gac gcg      721
85 Asp Glu Ser Ile Arg His Thr Leu Val Leu Arg Ser Leu Gln Asp Ala
86      225      230      235
88 tac aag gat cgt caa cgt atc ttt cgc cct ctc ccg ttg gca tgc cac      769
89 Tyr Lys Asp Arg Gln Arg Ile Phe Arg Pro Leu Pro Leu Ala Cys His
90      240      245      250
92 cgg ttg aag aac gcg ccg gaa tac gtc gaa tgg cac tca gca gaa aat      817
93 Arg Leu Lys Asn Ala Pro Glu Tyr Val Glu Trp His Ser Ala Glu Asn
94 255      260      265      270
96 ctt ttc cac tct atc tac aac gat gac aag cag aag aag ctc ttt acc      865
97 Leu Phe His Ser Ile Tyr Asn Asp Asp Lys Gln Lys Lys Leu Phe Thr
98      275      280      285
100 ctg ctg acg aac cat cgc tgc aca cga ctg gcg ctt acg ggc ggg tat      913
101 Leu Leu Thr Asn His Arg Cys Thr Arg Leu Ala Leu Thr Gly Gly Tyr
102      290      295      300
104 gag aag aag att ggc gct gcc gag gtc agg aat cta ctg gcc acc agg      961
105 Glu Lys Lys Ile Gly Ala Ala Glu Val Arg Asn Leu Leu Ala Thr Arg
106      305      310      315
108 aat cct agt tcg cag ctg gac agc tat atc atg gcg aag gta tat gta      1009
109 Asn Pro Ser Ser Gln Leu Asp Ser Tyr Ile Met Ala Lys Val Tyr Val
110      320      325      330
112 ctg gcg tcg gga gcg atc ggc aac cca cag att ctc tat aac tcg ggc      1057
113 Leu Ala Ser Gly Ala Ile Gly Asn Pro Gln Ile Leu Tyr Asn Ser Gly
114 335      340      345      350
116 ttc tct ggg cta cag gtc acg cca cgc aat gac tcg ttg atc ccc aac      1105
117 Phe Ser Gly Leu Gln Val Thr Pro Arg Asn Asp Ser Leu Ile Pro Asn
118      355      360      365
120 ctg ggg agg tac atc acg gag cag ccg atg gca ttt tgc cag ata gtc      1153
121 Leu Gly Arg Tyr Ile Thr Glu Gln Pro Met Ala Phe Cys Gln Ile Val
122      370      375      380
124 ttg agg cag gaa ttc gtc gac agc gtg cgc gac gat cct tat gga ctg      1201
125 Leu Arg Gln Glu Phe Val Asp Ser Val Arg Asp Asp Pro Tyr Gly Leu
126      385      390      395
128 cca tgg tgg aaa gaa gcc gtt gct caa cat att gcc aag aac ccg aca      1249
129 Pro Trp Trp Lys Glu Ala Val Ala Gln His Ile Ala Lys Asn Pro Thr
130      400      405      410
132 gat gca ctg ccc att ccg ttc cgc gat ccg gaa ccc cag gta aca acc      1297
133 Asp Ala Leu Pro Ile Pro Phe Arg Asp Pro Glu Pro Gln Val Thr Thr

```

## RAW SEQUENCE LISTING

DATE: 01/14/2002

PATENT APPLICATION: US/09/856,327

TIME: 10:50:58

Input Set : A:\0230-0157P.ST25.txt

Output Set: N:\CRF3\01142002\I856327.raw

```

134 415          420          425          430
136 cca ttt aca gaa gaa cac ccc tgg cac acg cag att cac cgc gat gct      1345
137 Pro Phe Thr Glu Glu His Pro Trp His Thr Gln Ile His Arg Asp Ala
138          435          440          445
140 ttt tcg tac ggt gcc gtc ggt cct gag gtg gac tct cgt gtc atc gtc      1393
141 Phe Ser Tyr Gly Ala Val Gly Pro Glu Val Asp Ser Arg Val Ile Val
142          450          455          460
144 gac ctg cgc tgg ttt ggc gca acc gac cct gaa gca aac aac ctt ttg      1441
145 Asp Leu Arg Trp Phe Gly Ala Thr Asp Pro Glu Ala Asn Asn Leu Leu
146          465          470          475
148 gtt ttc cag aac gat gtt caa gac ggg tac agt atg ccg cag ccg acg      1489
149 Val Phe Gln Asn Asp Val Gln Asp Gly Tyr Ser Met Pro Gln Pro Thr
150          480          485          490
152 ttc aga tat cga ccc agc act gcg tca aac gtg aga gca agg aaa atg      1537
153 Phe Arg Tyr Arg Pro Ser Thr Ala Ser Asn Val Arg Ala Arg Lys Met
154 495          500          505          510
156 atg gcc gat atg tgc gaa gtg gcg agc aac ttg gga ggt tat ttg ccc      1585
157 Met Ala Asp Met Cys Glu Val Ala Ser Asn Leu Gly Gly Tyr Leu Pro
158          515          520          525
160 acg tcc ccc ccg cag ttt atg gat cca ggc ctt gca ctt cat ctt gcg      1633
161 Thr Ser Pro Pro Gln Phe Met Asp Pro Gly Leu Ala Leu His Leu Ala
162          530          535          540
164 ggg act act cgc att ggc ttc gac aag gca act aca gtg gct gat aac      1681
165 Gly Thr Thr Arg Ile Gly Phe Asp Lys Ala Thr Thr Val Ala Asp Asn
166          545          550          555
168 aac tcg ctg gtc tgg gac ttt gcc aat ctt tat gtt gca ggc aat ggc      1729
169 Asn Ser Leu Val Trp Asp Phe Ala Asn Leu Tyr Val Ala Gly Asn Gly
170          560          565          570
172 acc atc agg acg ggc ttc ggc gag aac ccg aca ctt acg tcg atg tgc      1777
173 Thr Ile Arg Thr Gly Phe Gly Glu Asn Pro Thr Leu Thr Ser Met Cys
174 575          580          585          590
176 cac gct atc aag agc gcg agg agc atc atc aat aca ctc aag ggt ggg      1825
177 His Ala Ile Lys Ser Ala Arg Ser Ile Ile Asn Thr Leu Lys Gly Gly
178          595          600          605
180 act gac gga aaa aat aca ggc gag cat cgc aac ctt tgaggaagga      1871
181 Thr Asp Gly Lys Asn Thr Gly Glu His Arg Asn Leu
182          610          615
184 gcaacagcag tgtaaacaac cgcgtcaagt ggctacttca agttgaatgc attctggtcc      1931
186 cctaccatgt tgatgtgtac gataggcggt gaaagatttt gtgtattact gaacctgtac      1991
188 tttgtctgaa tagttatggc actatgattc atgttttaaaa aaaaaaaaaa aaaaaaaaaa      2051
190 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaa      2106
193 <210> SEQ ID NO: 2
194 <211> LENGTH: 618
195 <212> TYPE: PRT
196 <213> ORGANISM: Lyophyllum shimeji
198 <400> SEQUENCE: 2
200 Met Ser Leu Ser Thr Glu Gln Met Leu Arg Asp Tyr Pro Arg Ser Met
201 1          5          10          15
204 Gln Ile Asn Gly Gln Ile Pro Lys Asn Ala Ile His Glu Thr Tyr Gly

```

## RAW SEQUENCE LISTING

DATE: 01/14/2002

PATENT APPLICATION: US/09/856,327

TIME: 10:50:58

Input Set : A:\0230-0157P.ST25.txt

Output Set: N:\CRF3\01142002\I856327.raw

205		20		25		30	
208	Asn	Asp	Gly	Val	Asp	Val	Phe
209		35		40		45	
212	Thr	Tyr	Ala	Lys	Leu	Cys	Val
213		50		55		60	
216	Glu	Ile	Gly	Ala	Ala	Asp	Ser
217	65			70		75	
220	Thr	Ala	Val	Pro	Tyr	Val	Pro
221				85		90	
224	Phe	Gln	Lys	Asp	Ile	Asp	Arg
225				100		105	
228	Gln	Gln	Val	Ser	Val	Pro	Val
229				115		120	
232	Pro	Gly	Ala	Trp	Ser	Ala	Pro
233				130		135	
236	Lys	Asn	Pro	His	Gln	Arg	Glu
237	145					150	
240	Thr	Arg	Gly	Val	Gly	Met	Ser
241						165	
244	Arg	Ile	His	Pro	Pro	Met	Glu
245						180	
248	Leu	Ser	Asn	Asp	Pro	Ala	Glu
249						195	
252	Ser	Glu	Ala	Glu	Arg	Leu	Ile
253						210	
256	Ser	Ile	Arg	His	Thr	Leu	Val
257	225					230	
260	Asp	Arg	Gln	Arg	Ile	Phe	Arg
261						245	
264	Lys	Asn	Ala	Pro	Glu	Tyr	Val
265						260	
268	His	Ser	Ile	Tyr	Asn	Asp	Asp
269						275	
272	Thr	Asn	His	Arg	Cys	Thr	Arg
273						290	
276	Lys	Ile	Gly	Ala	Ala	Glu	Val
277	305					310	
280	Ser	Ser	Gln	Leu	Asp	Ser	Tyr
281						325	
284	Ser	Gly	Ala	Ile	Gly	Asn	Pro
285						340	
288	Gly	Leu	Gln	Val	Thr	Pro	Arg
289						355	
292	Arg	Tyr	Ile	Thr	Glu	Gln	Pro
293						370	
296	Gln	Glu	Phe	Val	Asp	Ser	Val
297	385					390	
300	Trp	Lys	Glu	Ala	Val	Ala	Gln
301						405	

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/856,327

DATE: 01/14/2002

TIME: 10:50:58


Input Set : A:\0230-0157P.ST25.txt

Output Set: N:\CRF3\01142002\I856327.raw

```

304 Leu Pro Ile Pro Phe Arg Asp Pro Glu Pro Gln Val Thr Thr Pro Phe
305           420           425           430
308 Thr Glu Glu His Pro Trp His Thr Gln Ile His Arg Asp Ala Phe Ser
309           435           440           445
312 Tyr Gly Ala Val Gly Pro Glu Val Asp Ser Arg Val Ile Val Asp Leu
313           450           455           460
316 Arg Trp Phe Gly Ala Thr Asp Pro Glu Ala Asn Asn Leu Leu Val Phe
317 465           470           475           480
320 Gln Asn Asp Val Gln Asp Gly Tyr Ser Met Pro Gln Pro Thr Phe Arg
321           485           490           495
324 Tyr Arg Pro Ser Thr Ala Ser Asn Val Arg Ala Arg Lys Met Met Ala
325           500           505           510
328 Asp Met Cys Glu Val Ala Ser Asn Leu Gly Gly Tyr Leu Pro Thr Ser
329           515           520           525
332 Pro Pro Gln Phe Met Asp Pro Gly Leu Ala Leu His Leu Ala Gly Thr
333           530           535           540
336 Thr Arg Ile Gly Phe Asp Lys Ala Thr Thr Val Ala Asp Asn Asn Ser
337 545           550           555           560
340 Leu Val Trp Asp Phe Ala Asn Leu Tyr Val Ala Gly Asn Gly Thr Ile
341           565           570           575
344 Arg Thr Gly Phe Gly Glu Asn Pro Thr Leu Thr Ser Met Cys His Ala
345           580           585           590
348 Ile Lys Ser Ala Arg Ser Ile Ile Asn Thr Leu Lys Gly Gly Thr Asp
349           595           600           605
352 Gly Lys Asn Thr Gly Glu His Arg Asn Leu
353           610           615
356 <210> SEQ ID NO: 3
357 <211> LENGTH: 30
358 <212> TYPE: PRT
359 <213> ORGANISM: Lyophyllum shimeji
361 <400> SEQUENCE: 3
363 Asn Ala Glu Glu Gly Thr Ala Val Pro Tyr Val Pro Gly Tyr His Lys
364 1           5           10           15
367 Lys Asn Glu Ile Glu Phe Gln Lys Asp Ile Asp Arg Phe Val
368           20           25           30
371 <210> SEQ ID NO: 4
372 <211> LENGTH: 24
373 <212> TYPE: PRT
374 <213> ORGANISM: Lyophyllum shimeji
376 <400> SEQUENCE: 4
378 Glu Phe Asp Glu Ser Ile Arg His Thr Leu Val Leu Arg Ser Leu Gln
379 1           5           10           15
382 Asp Ala Tyr Lys Asp Arg Gln Arg
383           20
386 <210> SEQ ID NO: 5
387 <211> LENGTH: 29
388 <212> TYPE: PRT
389 <213> ORGANISM: Lyophyllum shimeji
391 <400> SEQUENCE: 5

```

 Use of n and/or Xaa has been detected in the Sequence Listing.  
 Review the Sequence Listing to insure a corresponding  
 explanation is presented in the <220> to <223> fields of  
 each sequence using n or Xaa.

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/856,327

DATE: 01/14/2002

TIME: 10:50:59

Input Set : A:\0230-0157P.ST25.txt

Output Set: N:\CRF3\01142002\I856327.raw

L:435 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:465 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9  
L:483 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10  
L:501 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11  
L:519 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12